DOCUMENT RESUME

ED 322 175 TM 015 311

AUTHOR Renfrow, Donata D.; Kromrey, Jeffrey D.
TITLE Descriptions and Uses of Content-Specific Pedagogical Items.
PUB DATE Apr 90
NOTE 23p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Boston, MA, April 17-19, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Definitions; Elementary School Teachers; Elementary Secondary Education; Higher Education; *Knowledge Level; *Licensing Examinations (Professions); *Minimum Competency Testing; *Multiple Choice Tests; Secondary School Teachers; Teacher Education; *Test Construction; Test Items
IDENTIFIERS *Content Specific Test Items; *Teacher Competency Testing

ABSTRACT

After a brief review of recent research concerning teacher testing, this paper presents concrete examples of multiple-choice test items that assess more than teachers' minimal, basic knowledge. Specifically, multiple-choice test items used for the assessment of pedagogical content knowledge (C-P items) are contrasted with items used for the assessment of content knowledge (C items) and items used for the assessment of general pedagogical knowledge (P items). Following the contrasting examples, a working definition of C-P items is developed, practical considerations related to their development and use in testing programs are presented, and a categorization of the content of such items is proposed. The experiences of Florida and Georgia with regard to the development of subject-area certification tests for teachers are considered. The distinguishing features of C-P items are described. The following working definition of C-P items is proposed: "The class of C-P items includes those items for which the examinee's determination of the correct response depends upon knowledge of the treatment of content in educational situations." This definition excludes those items that solely address content, without an educational context, and those items that address general pedagogical principles in the absence of content-specific implications. Four major categories of items include: error diagnosis, communicating with the learner, organization of instruction, and learner characteristics. Two figures are included. (RLC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Donata Renfrow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED322175

DESCRIPTIONS AND USES OF CONTENT-SPECIFIC PEDAGOGICAL ITEMS

Donata D. Renfrow

Georgia State University

Jeffrey D. Kromrey

University of South Florida

A Paper Presented at the Annual Meeting of the National Council on Measurement in Education, Boston, April 17-19, 1990.

TM015311

BEST COPY AVAILABLE

Teacher testing has a long history, dating back to the 19th century. Shulman (1986) has provided an interesting look at the changing nature of teacher certification testing over the past century. Rudner (1988) summarized the types of tests used recently in state teacher certification programs. Most states involved in certification testing assess beginning teachers with performance-based evaluation, a paper-and-pencil test, or both. In addition, several states also use tests as a part of the recertification process. Some tests are designed to measure basic skills (reading, writing, and arithmetic); others are developed to measure basic pedagogical knowledge; and others purport to measure content area knowledge. Of states using content area knowledge tests, some (e.g., Connecticut, Florida, Georgia, Illinois, Oklahoma, Texas, West Virginia, and Washington, D.C.) use tests custom-designed for their own teacher population.

Teacher testing has not been popular with the teaching community. Court cases, such as the one in Georgia, have put teacher certification tests under considerable scrutiny (Jaeger & Bush, 1988). Challenges to existing paper-and-pencil tests have been based on both the format and the content of questions. These tests have been shown to have limited validity as predictors of teaching competence (Madaus & Pullin, 1987), and they have been shown to have a differential impact on minorities.

Another challenge on the basis of face validity may be minor from a psychometric perspective, but major from a political perspective. Content area test items typically address the content of the teacher certification fields (e.g., language arts, math, science, social studies) in a multiple-choice format. The tests, however, do not assess what teachers know about teaching those subjects to their students. Rudner (1988) has criticized current

teacher tests for being

> based on the logic that people who cannot pass a simple test of minimal, basic knowledge that is often acquired by eighth grade should not be placed in a position where they are responsible for the education of children. Such testing is a poor substitute for a valid test that measures the skills and attitudes needed to be a teacher.... (p. 19)

The need to consider more performance-based evaluations is clear. Several projects are underway to explore more "authentic" approaches to assessment (e.g., Leinhardt, 1990; Popham, 1989; Shulman, 1986, 1987, 1988).

Concurrent with the criticism of existing teacher testing programs, changes in the conception of teaching as a profession have been evident in recent years, charged with tremendous implications for teacher assessment (Shulman, 1987).

Shulman presented a categorization of the knowledge required for teaching: (a) content knowledge, (b) general pedagogical knowledge, (c) curriculum knowledge, (d) pedagogical content knowledge, (e) knowledge of the learner, (f) knowledge of educational contexts, and (g) knowledge of educational goals. Of these, Shulman asserted that pedagogical content knowledge may best delineate the knowledge base of teaching:

> But the key to distinguishing the knowledge base of teaching lies at the intersection of content and pedagogy, in the capacity of a teacher to transform the content knowledge he or she possesses into forms that are pedagogically powerful and yet adaptive to the variations in ability and background presented by the students. (Shulman, 1987, p. 15)

While the researchers continue the debate about how both the conception of teaching and the assessment of teaching should be changed, legislatures and

courts continue providing mandates for immediate action. Deadlines are set and states' departments of education are compelled to meet them, with or without the research community's resolution. For example, the Georgia Quality Basic Education Act describes the legislative mandate for teacher assessment:

> The State Board of Education shall require the applicant to demonstrate satisfactory proficiency on a test of specific subject matter or other professional knowledge appropriate to the applicant's field of specification.

During this transition period, at least four states are actively pursuing the assessment of content specific pedagogical knowledge, blending project timelines with the latest thinking in assessment. Much has been written about the Connecticut project in which a three-tier license exam for elementary educators is being prepared (e.g., Popham, 1988; DeLandshere & Mason, 1989; Carlson, 1989; Guiton & Delandshere, 1989; Potter & Sjoerdsma, 1988). Part of the Connecticut examination includes multiple-choice and free-response items for classroom videotapes. Texas, Florida, and Georgia, however, are approaching the measurement of content-specific pedagogy through the use of multiple-choice items.

Texas is considering content-specific pedagogy in its elementary education certification examination, but Florida and Georgia are approaching the issue of content-specific pedagogical questions in areas beyond elementary education. All of Florida's 44 content fields are developing items assessing content-specific pedagogy. Of the 13 fields completed in Georgia, ten are actively pursuing this new item type (e.g., K-12 Music, special education areas (six), and early and middle childhood tests).

The purpose of this paper is to present concrete examples of multiple-

choice test items that assess more than minimal, basic knowledge. Specifically, multiple-choice test items used for the assessment of pedagogical content knowledge (C-P items) are contrasted with items used for the assessment of content knowledge (C items) and items used for the assessment of general pedagogical knowledge (P items). Following the contrasting examples, a working definition of C-P test items is developed, practical considerations related to their development and use in testing programs are presented, and a categorization of the content of such items is suggested.

Two States' Experiences

During the development of subject-area certification tests for teachers, both Florida and Georgia have included C-P items. In the experience of both states, such items arose from the work of the subject-area test development and validation committees. By including such items in the examinations, the committees were not aware that they were producing something unique in the testing world. Rather, such items were necessary to measure certain of the skills toward which the test was directed.

Although some specifics differed, the general model followed in both Florida's and Georgia's subject-area test development included (a) drafting and validating lists of subject-area knowledge and skills, (b) conducting a job analysis survey of practicing teachers in the state to obtain ratings of the importance and job-relatedness of each skill statement, (c) modifying the list of knowledge and skills based upon the job analysis, (d) developing and validating item specifications, (e) writing, piloting, and validating

individual test items, (f) assembling and validating test forms. In Florida's experience, the need for C-P items originated in the drafting of the list of knowledge and skills. At this initial stage of test development, skills were identified that required the use of such items. This need was subsequently validated by the subject-area committees and further substantiated by the job analysis survey.

The discussion of the Georgia experience presented in this paper focuses primarily on the revision of the music education examination. The music educators' efforts to include content-specific pedagogical items resulted from both the Georgia job analysis and the negative reactions to the kinds of recall items (typically specific, whether low level or esoteric) that formed the previous certification examination. The educators sought to produce test questions that more closely approximated decisions and activities they encountered on a daily basis.

Although the focus of this paper is on the description of C-P items, it should be pointed out that educators in Georgia were also pursuing application-level, content-specific test items. In contrast to the single-fact, recall items (some trivial, some esoteric, but all easily verifiable and often written in 50 words or less), the educators sought to write items that approximated classroom situations more realistically -- items that required integrating information, analyzing, synthesizing, predicting, or comparing. Rather than measuring discrete building blocks, these items aimed at the measurement of ideas, concepts, and applications that were aggregates of the discrete pieces of information. The importance of this point will be apparent later in the paper. In evaluating certain items that reflect what teachers do with content in the classroom, the distinction between application-level

content items and C-P items loses its clarity.

C-P Items, C Items, and P Items

The remainder of this paper presents a description of the three types of examination items and presents the distinguishing features of the C-P item. Figure 1 presents a P item developed to measure the examinees' knowledge of presentation of concepts.[1] This item measures a general awareness of "what to do next" that may apply to any content field (Note that no content is presented in this generic pedagogy item.). This is contrasted with a C-P item from a test for certification in the field of Specific Learning Disabilities. This item also measures "what to do next," but the application is specifically embedded in teaching mathematics to a learning disabled student. An examinee's ability to answer this item requires knowledge of the mathematics content, blended with general pedagogical knowledge and the specific pedagogical techniques used in teaching the learning disabled.

Figure 2 presents two items from a subject-area test in Art. The C item measures knowledge of the use of media to achieve a desired art effect. In contrast, the C-P item measures the application of the use of media to a specific instructional setting (in this example, a first-grade class). Again, the C-P item measures a blend of content knowledge with pedagogical knowledge.

---

[1] The items in these figures were obtained from Florida study guides or item specifications from the content fields, not from operational test forms.

A Working Definition of C-P Items

To develop a working definition of C-P items, we began with Shulman's (1986) distinction that these items measure the knowledge and skill that distinguish the biology teacher from the biologist. As we examined representative items from several content areas, operational problems developed.

Our initial conceptual concern arose in fields such as elementary education and the various fields in exceptional child education. In such content areas, a separate discipline of practice, distinct from educative involvement, is not discernible. That is, in fields such as music or biology, one can readily visualize a distinction between the musician and the music teacher and between the biologist and the biology teacher. However, a non-instructional parallel profession for elementary education or teaching the emotionally handicapped does not exist. In the latter fields, it is necessary to distinguish between the many content areas that are taught and their corresponding noninstructional disciplines. Thus, when an elementary teacher is teaching life science, the C-P items relevant to the teaching assessment are those that distinguish the teaching of life science to primary-grades learners from the knowledge and skill used by the practicing biologist. When the same teacher is teaching arithmetic, the C-P items distinguish the teacher from the mathematician. In elementary education and exceptional child education, practice in many areas of content must be distinguished from teaching many areas of content.

A related concern is that many of these items can be answered by either

teachers of the content or practitioners of the content. The possession of the knowledge to answer the item is not the critical issue. Rather, it is the relationship between the knowledge assessed with the item and (a) the performance of the act of teaching the discipline, and (b) the practice of the discipline. The C-P item presented in Figure 1 may be readily answerable by mathematicians who have no experience in teaching the learning disabled. Similarly, the C-P item in Figure 2 may be answerable by commercial artists who have never taught art. The content knowledge of insightful practitioners combined with a general familiarity with children and learning will probably yield a greater than chance probability of correct responses to these items from nonteacher examinees. These items, however, reflect what teachers of these fields do in the process of teaching, not what mathematicians or commercial artists do as they engage in their respective disciplines.

A third point of clarification is the need to avoid cosmetic content-pedagogy. If the P item in Figure 1 was revised so that the stem asked for the best method for teaching the concept of "greater than," the item would superficially appear to be a C-P item. Such a change may increase the face validity of the item for a subject-area test in mathematics, but the knowledge assessed by the item remains that of general pedagogy. The mere insertion of content-specific references does not change a P item to a C-P item. In the C-P items, the content aspects and the pedagogical aspects are interwoven. Removing the content from a C-P item or removing the pedagogy from a C-P item essentially destroys the item's intent and changes the nature of the knowledge being measured. Further, because of the subject-area specificity, the items are not interchangeable across content fields. For example, changing the content area of the C-P item in Figure 1 from teaching addition to teaching

letter recognition would require a complete rewrite of the item.

With these three initial clarifications of C-P items, the following working definition is proposed:

> The class of C-P items includes those items for which the examinee's determination of the correct response depends upon knowledge of the treatment of content in educational situations.

This definition excludes those items that solely address content, without an educational context, and those items that address general pedagogical principles in the absence of content-specific implications. Many examples of C-P items suggest a metacognitive component. The examinee must evaluate instructional aspects of the content and recognize logical errors in application without actually being in the classroom experiencing the events. Shulman (1987) has pointed out that "practitioners know a great deal that they have never even tried to articulate" (p. 12). Because of this lack of teachers' experience in reflecting upon their teaching, these items may be more difficult than application items that measure only content or only pedagogy.

Practical Issues

Carlson (1989) pointed out that C-P items are typically more difficult to write than either C items or P items. The experiences in Florida and Georgia substantiate this increased level of difficulty. Test blueprints are often

vague in detailing this type of item. Most educators who are writing items for these tests have only experienced the diagnostic testing model for producing questions. While it is relatively easy to produce single-factual recall items, most writers find it exceptionally difficult to produce questions measuring what they do from the integrated C-P perspective.

These items call for a metacognitive awareness of the teaching process. The production of an item requires the writer to do a "freeze frame" on a teaching situation and to call up to conscious awareness all of the elements and relationships that impact a final decision (e.g., to select or reject a particular choral score, to simplify a complex set of rhythms for the clarinets, to plan rehearsal activities for a particular band piece). Veteran teachers working as item writers often work intuitively and so quickly that it is difficult for them to be conscious of the variables that impact their decisions. It is also difficult for them to recall instances of classroom scenes that would make good scenarios for test items. Although the process of item writing is difficult when writers are required to sample important content, the process takes on additional complexity when this metacognitive analysis is also required and the content and pedagogy become interwoven.

A frustration for the trainers of the item writers has been in developing the kinds of examples and questioning strategy that guide the item writers through this process with a little less tension and frustration. Both authors agree that our recent experiences will make the next training session much easier. This paper summarizes examples and strategies that evolved during the training as we were learning what to do. Court deadlines and real dollar budgets typically necessitate learning as you go.

Carlson (1989) has already discussed the difficulties in developing items

with one (and only one) correct option. Each item's stimulus, considered in isolation, suggests many correct responses. The item writer's dilemma is to provide only one of these correct responses and three plausible options that are "less correct" than the key. Obtaining consensus on the correct option is often difficult. In an effort to maintain the plausibility of the foils, the distinction between a correct option and one that is "almost good enough" becomes fuzzy. From a communication theory standpoint, the difficulty is easy to explain. Because we are trying to capture an actual "frame" from the classroom "reel," we are limited by the words to describe the picture. Word choices are based on our individual hidden -- or not hidden -- biases. We are now grappling with where our individual attention is focused, the interpretations we put on the "symbols" we are focusing on, and their relationship to the rest of the picture.

The classic picture of the "lady," familiar from the psychology of perception, shows the dilemma quite well. Focusing on one set of features shows the image to be one of an attractive young woman, while a change in perspective shows the same image to be one of an old woman with a wart on her nose. The item writers must scope the whole "frame" to avoid ignoring other critical features that may be attended to by the examinees.

In addition to being difficult to write and difficult to assure the provision of a single correct option, these items are more difficult to edit than simpler item types. Field testing is an invaluable aid in the editorial and revision process, but most states do not have the luxury to field test the items adequately. Georgia's music items were field-tested by being imbedded within the regular operational test forms. Some items were able to be field-tested a second time when data on distractors suggested weaknesses or flaws.

Other states' field-test procedures for new items provide only small samples of students in teacher education programs, students who, in some cases (e.g., Connecticut), were paid to participate in the pilot.

Issues with the statistical properties of these items are addressed elsewhere (Delandshere & Guiton, 1990; Renfrow et al, 1990), but because this type of item is a recent development, research on the psychometric performance of these items is just beginning.

A Proposed Categorization of C-P Items

Based upon a review of C-P items developed for a variety of subject-area tests, four major categories of items have been identified.[2] By the very nature of its development, this list of categories must be viewed as incomplete. The C-P items reviewed in the process of developing this framework were not written to determine the number of different ways such items can appear, but were written to measure specific skills identified as important for inclusion on teacher subject-area certification examinations. As more states gain experience in developing this type of item, it is anticipated that new categories and new variations on these four categories will appear.

A second problem is that the proposed categories are not mutually

---

[2] Previous efforts have been directed at the development of a framework for teacher knowledge (e.g., Shulman, 1987; Tamir, 1988; Smith & Neale, 1989) but, aside from Carlson's (1989) list of item types, this is probably the first attempt to develop a framework to classify test items designed to measure content-specific pedagogical knowledge.

exclusive. A single C-P item may contain aspects of several of the following categories.

Despite the incompleteness and overlap in the categories, it is hoped that the delineation of these categories will assist other test developers in writing effective C-P items.

Category 1: Error Diagnosis

One of the most commonly occurring categories of C-P items is that which includes an error-diagnosis component. The stimulus presents an example of student work (e.g., 10 measures of a musical score are played and the measures are printed; a student's solution to a series of mathematical problems is presented; several paragraphs of text with oral reading errors are marked) and the examinee is required to respond to the example. Problem analysis can occur in several different ways:

- Identify the error (e.g., violins played in C natural instead of in C sharp).
- Identify the student's logical error (task analysis) either by naming the problem or replicating it (e.g., student did not convert to a common denominator; student misinterpreted 2/2 time and played half notes as two beats instead of as one beat).

Category 2: Communicating with the Learner

The second major category of C-P items is that which deals with appropriate communications with learners. These types of items appear in the following major ways:

- Evaluate student homework (e.g., which feedback is most appropriate for a six-year-old first grader who wrote a story about "nites in shng armr ftng dragnz"; a student's customer letter responding to a delayed order for a business communication class).

- Simulate a dialogue between teacher and student(s) as the item stimulus, to show student confusions. The response required is a "next step" activity or query that would best lead the student(s) to understanding the problem and resolving the confusion.

Category 3: Organization of Instruction

This category of test items includes those which focus on teacher plans for instruction rather than on students' errors or the nature of teacher-student communication. For instance, the item stimulus may describe a group of students and an instructional objective. The item response options would be teaching activities, one of which is most appropriate for the group and the objective. Variations on this basic item type are:

- An activity is described that did not result in successful instruction and the item asks for an alternate activity.

- A failed activity is described, and the respondent provides a plausible reason for the failure.

- An activity is described, some part of which is inappropriate. The respondent identifies how the activity can be corrected or why it is inappropriate.

- A failed activity is described and a successful corrected activity is described. The respondent identifies a reason the correction worked.

- Given a set of classroom resources that are available, the item asks for a plausible activity or how to compensate for a limit in the resources (e.g., the first violinist contracted mono two days before the concert. How would you compensate for this absence? A chemistry teacher is out of compound X, but has plenty of U, V, and W. What if any of these can be substituted to complete the planned lesson?).

Further variations on this theme include items that describe a group of students and:

- Given an activity, identify the objective.
- Order a set of activities or skills in the most appropriate manner.
- Translate material to a different level (e.g., the trumpets can't play this part. How should it be simplified for rehearsal?).

Another subcategory of instructional organization questions is that which addresses content-specific methods, materials, and evaluation:

- Content-specific methods and materials (e.g., the stimulus presents a musical score and queries about its use with a particular group of students; the stimulus presents a description of a reading activity which the examinee must identify as one particular method of teaching reading).
- Questions on formal and informal evaluations in the content area include the selection of an appropriate evaluation, interpretation of the results, the drawing of reasonable conclusions from evaluation results, and the prediction of appropriate instructional directions and next steps.

Category 4: Learner Characteristics

The final category of C-P items is that which includes items addressing the examinees' knowledge of developmental norms within the content area or the expected sequences of skill development and the progression of competence in the discipline (e.g., a teacher is having trouble teaching addition of fractions to first graders. Why?).

Conclusions

This paper extends Shulman's ideas and operationalizes them by describing how two states have begun developing C-P items. The categorization proposed has been elaborated with sample items, to a greater extent than has been presented previously.

Two issues have emerged during the preparation of this paper. As presented in the development of the working definition of C-P items, we have begun to see some implications of the distinction between the teaching specialties that have "doers" outside the school setting (e.g., chemists, historians, journalists, literary critics, musicians, mathematicians) and those that don't (e.g., early and middle childhood teachers, reading specialists, special education teachers). In constructing valid content-area tests for these fields, more C-P items would be likely to appear on a test for the latter group, where most of the content knowledge is couched in the teaching framework. The other content fields, however, lend themselves to content items that reflect higher-order thinking skills, that are content specific and could be answered by either a member of that profession or by a teacher in that field. In tests for these areas, C-P items would represent only a small proportion of a valid examination.

This point leads us to our second issue. These C-P items, as we have currently categorized them, have not been used to any great extent on previously developed content-area tests, where lower-level questions have predominated. Musicians could easily answer the error detection items, and mathematicians likely could identify the mistaken operations in an algebraic calculation. Musicians could just as easily complete the last four measures of

a score (selecting which option best completes the song in its style) as a teacher could, and yet such a skill is essential for helping students learn to write their own music.

As we have considered Shulman's distinction, trying to isolate knowledge that is specific to the teacher, we found that if we adopted his premise completely, items such as the error detection items would need to be eliminated from the C-P category. Since we feel that error analysis is integral to teaching -- seeing how students understand the subject -- we feel it should be included in this category, even if nonteaching professionals could answer the question as well.

The development of C-P items presents unique demands on test developers. Training item writers is more difficult; more time must be allocated for item writing and review; and several cycles of piloting and revision of items may be required to produce an item of quality. We have found, however, that we can construct multiple-choice items that lead to consensus on the correct response. More importantly, teachers involved in the projects agree that these items accurately reflect the process of teaching the content.

References

Carlson, R.E. (1989). The intersection of subject matter and pedagogical knowledge in teacher assessment: Practical considerations for item development. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Delandshere, G., & Guiton, G. (1990). Content-specific assessment -- measurement concerns. Paper presented at the meeting of the National Council on Measurement in Education, Boston, MA.

Delandshere, G., & Mason, T.C. (1989). The effect of opportunity to learn on test item statistics. Paper presented at the meeting of the National Council on Measurement in Education, San Francisco, CA.

Guiton, G.W., & Delandshere, G. (1989). Making sense of field-test item data: An experiment involving an innovative teacher certification test Paper presented at the meeting of the National Council on Measurement in Education, San Francisco, CA.

Leinhardt, G. (1990). Capturing craft knowledge in teaching. Educational Researcher, 19, 18-25.

Madaus, G.F., & Pullin, D. (1987). Teacher certification tests: Do they really measure what we need to know? Phi Delta Kappan, 31-38.

Millman, J. (1988). Review of test development and score reporting procedures. In R. Jaeger & J.C. Bush (Eds.), An Evaluation of the Georgia Teacher Certification Testing Program.

Popham, W.J. (1988). The field-testing of items: Traditionalism's trap for innovative teacher assessment. Paper presented at the Directors of State Testing Programs, New York.

Potter, L., & Sjoerdsma, R. (1988). Elementary education licensure examinations for the '90s. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Renfrow, D.D., Gabrielson, S., Spaner, S., & Cramer, S.E. (1990). Multiple-choice items on a teacher certification test: Can they be used to assess content-specific pedagogy. Paper presented at the meeting of the National Council on Measurement in Education, Boston, MA.

Reynolds, Anne (1989). Personal communication.

Rudner, L. (1988). Teacher testing -- an update. Education Measurement: Issues and Practice, 7, 16-19.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15, 4-14.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 54, 1-22.

Shulman, L.S. (1987). Assessment for teaching: An initiation for the profession. Phi Delta Kappan, September, 38-44.

Shulman, L.S. (1988). A union of insufficiencies: Strategies for teacher assessment in a period of educational reform. Educational Leadership, 36-41.

Smith, D.C., & Neale, D.C. (1989). The construction of subject matter knowledge in primary science teaching. Teaching and Teacher Education, 5, 1-20.

Tamir, P. (1988). Subject matter and related pedagogical knowledge in teacher education. Teaching and Teacher Education, 4, 99-110.

Figure 1

Items to Measure the Presentation of Concepts

P item	C-P item
According to current research, the most effective method for teaching concepts is to provide A. definition, examples, and non-examples. B. verbal drill and practice. C. visual, auditory, and kinesthetic activities. D. work sheets for written practice.	Mrs. Stevens will introduce addition to her first-grade SLD class. The best hierarchy for her to follow is to have the students A. recognize the words addend and sum; understand the "+" sign; compute sums less than ten; understand place value concerning regrouping tens and ones. B. estimate sums; understand the "+" sign; understand place value of ones and tens; compute sums less than ten. C. find missing addends; understand place value of ones and tens; understand the "+" sign; understand place value concerning regroupings of tens and ones. D. recognize the words addend and sum; estimate sums; understand place value of ones and tens; compute sums less than ten.

Figure 2

Measurement of Content Knowledge and Pedagogical Content Knowledge

C item	C-P item
An artist drawing illustrations for a book with a somber mood would most likely use A. pen and ink washes. B. pastels, wet and dry technique. C. thick and thin markers. D. colored pencils and watercolor washes.	To introduce gesture drawing to a class of first-grade students, the best material is A. crayon. B. vine charcoal. C. oil pastels. D. India ink.